

DOI: 10.19650/j.cnki.cjsi.J1804427

基于卷积神经网络的汉字编码标记点检测识别*

陶 聪, 施 云, 张丽艳

(南京航空航天大学机电学院 南京 210016)

摘 要:近景摄影测量中采用的标记点要求具有唯一身份号并能在图像中被精确识别定位。设计了一种以汉字作为编码特征的编码标记点,提出了一种基于卷积神经网络的编码标记点检测识别方法。首先采用基于相机成像原理的虚拟相机法,自动生成大量汉字编码点模拟图像作为训练样本,并据此训练卷积神经网络成为汉字编码点识别网络。根据一系列编码点筛选准则分割得到实拍汉字编码点,然后用编码点识别网络对其身份号进行识别,最后通过中心定位算法定位编码点中心。实验结果表明构建的识别网络对汉字编码点识别率可达 97.67%,且受噪声、投影角度、图像对比度、亮度等因素的影响小;分割算法鲁棒性强,能准确分割出汉字编码点;中心定位算法对编码点中心的定位精度高。

关键词:近景摄影测量;汉字编码标记点;卷积神经网络;图像分割;图像识别;中心定位

中图分类号: TP391.41 **文献标识码:** A **国家标准学科分类代码:** 520.2040

Detection and recognition of Chinese character coded marks based on convolutional neural network

Tao Cong, Shi Yun, Zhang Liyan

(College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: In close-range photogrammetry, it is required that the adopted coded marks must have unique identification number and can be identified as well as located accurately in the image. In this paper, a kind of coded marks with Chinese character as encoding characteristic is designed, and a detection recognition method for the coded marks is proposed based on convolutional neural network. Firstly, a virtual camera method based on the camera imaging principle is used to automatically generate large amount of simulative images of the designed Chinese character coded marks, which are used as training samples. These samples are used to train the convolutional neural network that is used as the recognition network of Chinese character coded marks. The real captured Chinese character coded marks in the measurement images are detected with a series of cede mark sifting criteria, and the identification number is identified with the coded mark recognition network. Finally, the ceded mark centers are located with the center location algorithm. The experiment results show that the proposed recognition network has strong ability for recognizing the Chinese character coded marks, the recognition rate reaches 97.67%. The proposed method is less affected by noise, projection angle, image contrast and brightness changes, and possesses strong robustness. The proposed method can accurately segment the Chinese character coded marks and the center location algorithm can accurately locate the mark centers.

Keywords: close-range photogrammetry; Chinese character coded mark; convolutional neural network; image segmentation; image recognition; central positioning

0 引 言

国内外现有的编码标记点主要有环状^[1]、圆形^[2]、正方形^[3]、点状^[4]、分布式^[5]、带颜色信息^[6]等编码方

式。环状、圆形及正方形等编码方式的识别算法简单快速稳定,但其编码受旋转影响,属于相对编码方式,编码空间较小,编码数量有限。点状和分布式编码方式则引入了设计坐标系,不受旋转影响,属于绝对编码方式,编码空间较大,编码数量比相对编码方式多,但其识别算

法易受噪声及变形影响,且由于识别算法应用了模式匹配原理,识别过程较为复杂且耗时。带颜色信息的编码方式其识别算法则容易受光照等因素影响。二维码^[7]如今得到广泛使用,但其抗模糊性较差,在成像模糊效应下识别率会大大降低,因此不适合作为标记点编码应用于近景摄影测量中。

由此可见,编码方式的复杂度受限于识别算法的鲁棒性,若要保证识别算法的鲁棒性,则编码一般只能在比较有限的编码空间内进行,这不仅限制了编码数量,在实际应用中,还容易受到光照、成像质量等影响,导致标记点识别错误而产生误匹配,最终影响三维重建;而若要扩大编码空间,提升编码复杂度以增加编码数量,则鲁棒性就难以保证,且识别过程中所消耗时间也会增加。因此,如何在扩大编码空间,提升编码复杂度,增加编码数量的同时,保证该编码能被算法快速准确地识别,仍是值得深入研究的问题。

近年来,深度学习已成为机器视觉的一个重要研究领域,其中卷积神经网络在分类、识别方面更有突破性进展。例如,将神经网络应用于传动部位故障分类^[8]与高光谱图像分类^[9];在 ImageNet 2017 比赛中获得冠军的 SENet^[10];在目标检测领域获取成功的基于区域特征提取的 R-CNN^[11]等等。卷积神经网络无需人为设计算法提取图像特征,卷积层可自动完成图像特征提取,成功避免了因编码复杂度上升而导致的识别算法复杂度上升问题。同时由于硬件设备的飞速发展,高性能的 GPU 使各种对计算能力要求极高的神经网络同样得以实现应用。

本文设计了一种以汉字作为编码特征的编码标记点,提出了一种基于卷积神经网络的识别网络用于识别该种编码点。汉字的数量保证了编码标记点足够大的编码空间,汉字在一定模糊效应下仍能具有基本架构特征保证了一定的抗模糊能力,基于卷积神经网络的识别网络则保证了识别汉字编码点的稳定性及效率。为了解决难以获取足量卷积神经网络训练样本的问题,本文采用了虚拟相机法并提出了干扰模型,两者结合可方便地模拟生成足量且优质的图像用于训练网络。本文还提出了汉字编码标记点的分割算法及中心定位算法。实验证明,本文通过虚拟相机法生成的模拟图像对卷积神经网络训练效果较好,且该网络识别汉字编码标记点时受光照、成像质量等因素影响小,识别稳定性好、效率高,分割算法鲁棒性好,中心定位算法精度高。

1 汉字编码点设计

1.1 汉字编码点结构

本文提出的基于卷积神经网络的编码标记点检测与识别方法不受特定编码特征识别算法限制,只要是具有

唯一性的特征都可用于编码标记点的编码设计。汉字作为一种文字符号,每个汉字本身都不同于其他汉字,具有唯一性的特征,且常用汉字数高达 7 000 个,汉字总量更是高达 90 000 个,使用汉字作为编码标记点的特征,具有很大的编码空间,能满足标记点的种类数量要求。由于存在现成的汉字字形库可以利用,非常便于挑选产生特征明显的大量汉字编码点,本文采用汉字作为编码标记点的特征,设计出一种汉字编码标记点(汉字编码点)。

图 1 所示为本文提出的汉字编码点的一个示例,汉字编码点的整体形状为正方形,其编码区域由黑色背景与白色编码汉字“赵”构成,定位区域由白色辅助圆内嵌黑色定位圆组成,且两圆为同心圆,一般取方形编码点正中心为圆心,近景摄影测量中即以中央黑色定位圆的圆心作为三维重建的定位点,白色辅助圆则用于中心定位中评价黑色定位圆圆心的定位精度。为了使汉字编码点的内部定位区域与外部编码区域有明显界限,在两者之间添加了一个以定位点为圆心的黑色大圆作为隔离。本文使用上述标记点时具体尺寸设定已在图 1 标出,该尺寸是根据本文实验时拍摄场景大小所调整的,在不同应用中,可对标记点的尺寸进行等比缩放以适应拍摄场景大小。

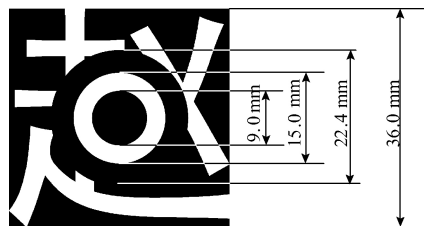


图 1 汉字编码标记点

Fig.1 Chinese character coded mark

1.2 编码汉字选择原则

有些汉字由于某种局限性不适合作为编码汉字,本文总结了两条编码汉字筛选原则以剔除这些汉字。

1) 无旋转歧义原则:不使用经过旋转后非常相似,极易引起混淆的汉字作为编码汉字。

图 2 所示为汉字“干”、“土”、“丫”、“人”作为编码汉字的汉字编码点,显然汉字“干”与“土”,汉字“丫”与“人”经过旋转后非常容易混淆,因此避免使用此类具有旋转相似性的汉字作为编码汉字。

2) 无遮盖歧义原则:不使用覆盖了辅助圆、定位圆后非常相似,极易引起混淆的汉字作为编码汉字。

图 3 所示为汉字“凡”、“几”、“九”、“丸”覆盖了辅助圆、定位圆后生成的汉字编码点,显然汉字“凡”与“几”,汉字“九”与“丸”经过编码点化后非常容易混淆,因此避免使用此类区分特征仅在中心部位,经编码点化

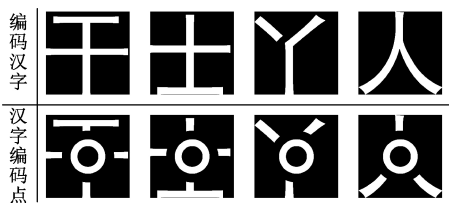


图2 旋转歧义的汉字编码点

Fig.2 Rotational ambiguous Chinese character coded marks

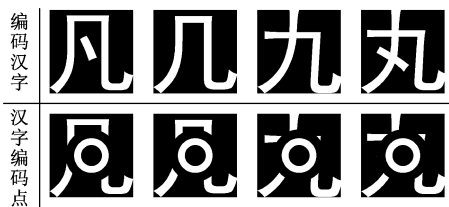


图3 遮盖歧义的汉字编码点

Fig.3 Covering ambiguous Chinese character coded marks

后难以区分的汉字作为编码汉字。

2 汉字编码点识别网络设计及训练

2.1 虚拟相机法生成训练样本

将卷积神经网络训练成能准确识别汉字编码点的识别网络,需投入大量不同汉字编码点在不同姿态下的成像图片进行训练,若采用实际拍摄方式获取图像,则效率极其低下且难以保证训练样本中汉字编码点成像能包含足够多的视角及姿态。因此,本文采用了一种基于相机成像原理的虚拟相机法,该方法可模拟生成各种汉字编码点在不同姿态下的成像图像(模拟图像),以此获取足量训练样本保证卷积神经网络的成功训练。

图4所示为虚拟相机法成像原理,该虚拟相机模型将汉字编码点从编码点坐标系 $O_{obj}X_{obj}Y_{obj}Z_{obj}$ 透视变换到像素坐标系 O_2uv ,其过程如下

$$s[u \ v \ 1]^T = K[R \ t] [X_{obj} \ Y_{obj} \ Z_{obj} \ 1]^T \quad (1)$$

其中,相机内参矩阵 K 用于描述相机坐标系 $O_cX_cY_cZ_c$ 与像素坐标系 O_2uv 之间的关系,可通过相机标定得到, $[X_{obj} \ Y_{obj} \ Z_{obj}]^T$ 为汉字编码点上任意点 D 的坐标,旋转矩阵 R 和平移向量 t 可用于描述点 D 从编码点坐标系 $O_{obj}X_{obj}Y_{obj}Z_{obj}$ 下的坐标到相机坐标系 $O_cX_cY_cZ_c$ 下的坐标的变换过程, s 为尺度因子。为了生成汉字编码点在不同姿态下的成像图片,需设定不同的 R, t , 用于模拟虚拟相机以不同角度拍摄汉字编码点。本文通过欧拉角设定汉字编码点空间旋转,将其表示为向量形式即 $r = [r_x, r_y, r_z]^T$ (姿态向量),其中 r_x, r_y, r_z 分别为编码点坐标系 $O_{obj}X_{obj}Y_{obj}Z_{obj}$ 依次绕自身轴 $O_{obj}X_{obj}, O_{obj}Y_{obj}, O_{obj}Z_{obj}$ 的

旋转角度,通过相关公式^[12]可将 r 转换为 R 。本方法中编码点坐标系 $O_{obj}X_{obj}Y_{obj}Z_{obj}$ 到相机坐标系 $O_cX_cY_cZ_c$ 的位移固定且不作平移改变,因此 t 为固定值。

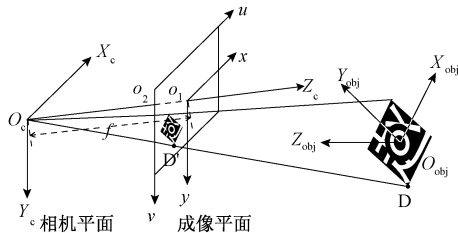


图4 虚拟相机法成像原理

Fig.4 The imaging of virtual camera method

实际测量中若相机光轴与编码点法矢所在直线所夹锐角(投影角度)过大,编码点透视成像的变形程度会比较严重。为了保证编码点识别及定位精度,应尽量避免投影角度过大。因此,使用虚拟相机法模拟生成训练样本时,将编码点坐标系 $O_{obj}X_{obj}Y_{obj}Z_{obj}$ 的姿态向量 $r = [r_x, r_y, r_z]^T$ 中各变量设定范围如下

$$r_x, r_y \in \left[-\frac{\pi}{3}, \frac{\pi}{3}\right], r_z \in [-\pi, \pi] \quad (2)$$

综上所述,虚拟相机法只需预先设定虚拟相机内参数矩阵 K , 平移向量 t , 即可通过随机设定编码点坐标系 $O_{obj}X_{obj}Y_{obj}Z_{obj}$ 的姿态向量 $r = [r_x, r_y, r_z]^T$, 将三维编码点坐标系 $O_{obj}X_{obj}Y_{obj}Z_{obj}$ 中的汉字编码点经过透视变换到二维的像素坐标系 O_2uv 中, 模拟生成汉字编码点在不同姿态下的成像图片。图5所示为以汉字“赵”作为编码特征的汉字编码点在不同姿态下的模拟图像。



图5 汉字编码点模拟图像

Fig.5 Simulative images of Chinese character coded marks

2.2 汉字编码点识别网络设计

本文设计的汉字编码点识别网络结构如图6所示,其中卷积层用于提取汉字编码点特征并生成特征图;非线性激活函数 $\text{ReLU}^{[13]}$ 可为识别网络增加非线性因素;最大值池化可对特征图进行降维处理以便后续进一步提取更抽象的图像特征^[14];随机神经元失活可防止神经网络出现过拟合问题^[15],系数0.25表示失活的神经元比例为25%;扁平处理目的是将原特征图从张量形式降维成可被全连接层接收的向量形式;全连接层的系数表示全连接层中向量的维数;Softmax 是分类器,可输出最终

分类结果。将汉字编码点输入识别网络进行分类后,最终可得到其类别,完成编码点识别。由图6中可视化图像可知,卷积网络的层数越深,所提取的特征图像越抽

象。卷积网络层提取的特征会具有很强的判别能力以及泛化性^[16],因此本文用了4个卷积网络层以保证汉字编码点识别网络的性能。

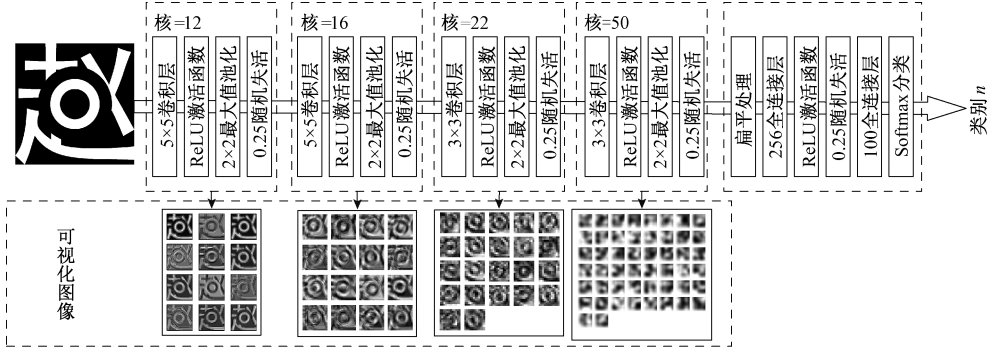


图6 汉字编码点识别网络结构及可视化图

Fig.6 Chinese character coded mark recognition network structure and visualized images

2.3 汉字编码点识别网络训练

卷积神经网络的训练目标是将损失函数最小化,本文的汉字编码点识别网络使用的损失函数为交叉熵损失函数^[17] $L(W, b)$ 。设训练样本共有 N 类编码点,某汉字编码点实际类别为 n ,将实际输出(概率分布)定义为 $q = [q_1, q_2, \dots, q_i, \dots, q_N]^T$,期望输出(概率分布)定义为 $p = [p_1, p_2, \dots, p_i, \dots, p_N]^T, i \in [1, N]$,则

$$q_i = \frac{\exp(W_i^T x + b_i)}{\sum_{j=1}^N \exp(W_j^T x + b_j)} p_i = \begin{cases} 1 & i = n \\ 0 & i \neq n \end{cases} \quad (3)$$

式中: x 为当前网络层的输入; W 为当前网络层的权重; b 为当前网络层的偏置。交叉熵函数 $L(W, b)$ 则刻画了实际输出(概率分布) $q = [q_1, q_2, \dots, q_i, \dots, q_N]^T$ 与期望输出(概率分布) $p = [p_1, p_2, \dots, p_i, \dots, p_N]^T$ 的距离,即:

$$L(W, b) = - \sum_{i=1}^N p_i \log q_i \quad (4)$$

本文采用小批量随机梯度下降法^[18]作为该识别网络的优化算法,其中批量数为128,用反向传播法^[19]对神经网络各层参数(W 和 b)进行更新,初始学习率 η 设为0.1,用于控制网络层参数更新速率,每次更新后都将学习率 η 降低 10^{-6} 以减小网络层参数的更新速率。为了防止训练神经网络时出现过拟合问题,本文使用 L_2 范数来控制网络层参数过拟合,并用参数 λ 控制 L_2 范数防止过拟合的强度。网络层参数更新过程可表示为:

$$W'_i = W_i - \eta \frac{\partial \left(L(W, b) + \frac{\lambda}{2} W^T W \right)}{\partial W_i} \quad (5)$$

$$b'_i = b_i - \eta \frac{\partial \left(L(W, b) + \frac{\lambda}{2} W^T W \right)}{\partial b_i} \quad (6)$$

本文先随机挑选了100种汉字编码点,并用虚拟相机法生成36万份图像样本,其尺寸均为 68×68 像素,该图像样本包含汉字编码点在各种不同姿态下的成像,具体成像姿态以编码点坐标系姿态向量定义(详见2.1节),然后将模拟图像样本分为训练样本和测试样本,其中训练样本为32万份,测试样本为4万份。实验时,分别在训练样本中随机抽取4万份、8万份、12万份、16万份及32万份训练样本对汉字编码点识别网络进行独立训练,一共可训练得到5个独立的识别模型。训练汉字编码点识别网络的过程中,使用型号为NVIDIA GeForce GTX 960显卡进行并行计算加速。每个识别模型在每个训练周期结束后都会得到该次训练的损失函数值(详见式(4)),当损失函数值开始趋于稳定不再明显减小时,则认为该识别模型已经训练完毕。当上述5个汉字编码点识别网络训练完毕后,用模拟生成的4万份测试样本分别对其进行识别率测试,结果如图7所示。

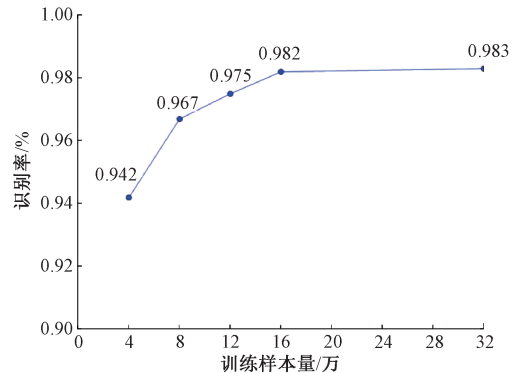


图7 训练样本量与网络识别率关系折线图

Fig.7 Relationship line chart of training sample size vs. network recognition rate

由图 7 可知,汉字编码点识别网络对汉字编码点的识别率随训练样本量的增加而上升,但当训练样本量到达 16 万时,其识别率已达 98.2% 且上升率非常小,训练样本量为 32 万时识别率仅上升 0.1%,因此本文将训练样本量设定为 16 万,并以当训练样本量为 16 万时损失函数趋于稳定的训练周期 50 作为汉字编码点识别网络的训练周期。

3 实拍汉字编码点检测

3.1 汉字编码点分割

对实拍图像中的汉字编码点进行识别和定位时,首先需要将布设于待测物体表面的汉字编码点从周围背景中分割出来,这是后续利用汉字编码点识别网络对其进行识别的前提。本文对汉字编码点的分割过程如图 8 所示,具体可分为 6 个步骤。

1) 输入包含待分割汉字编码点的灰度图像,图 8(a)所示为布设了汉字编码点的 LOGO 纸板,其中 LOGO 纸板上的汉字编码点即为待分割的目标。

2) 对输入的灰度图像进行形态学处理,即先二值化再形态学闭运算。形态学闭运算的目的是消除每个待分割汉字编码点内部阻碍连通的黑色区域,使其白色部分融合为一整个区域,则一个白色区域即为一个汉字编码点所在区域,处理后如图 8(b)所示。

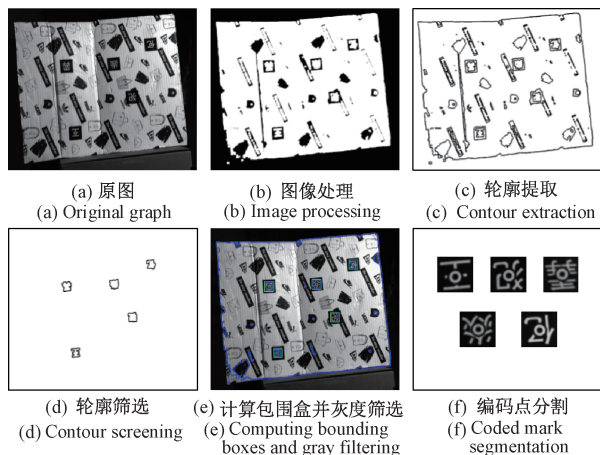


图 8 汉字编码点分割

Fig.8 Chinese character coded mark segmentation

3) 利用 Canny^[20]算子从形态学处理后的二值图像中提取轮廓,如图 8(c)所示,步骤 2)得到的每个包含汉字编码点的白色区域都被提取为一个连续轮廓。

4) 利用轮廓筛选准则筛选步骤 3) 获取的所有轮廓,获取包含待提取汉字编码点的轮廓,如图 8(d)所示。轮廓筛选准则有两条:

(1) 面积周长准则:根据轮廓的面积和周长进行

筛选;

(2) 最小外接矩形宽高比准则:根据轮廓的最小外接矩形宽高比进行筛选;

5) 计算筛选后轮廓的最小包围盒,根据最小包围盒参数在输入图像中截取图像区域,如图 8(e)所示,并对截取的图像进行二值化处理,然后计算该区域内的灰度均值,设定灰度均值阈值,若区域内图像灰度均值在阈值范围内则保留该区域,否则剔除;

6) 将经灰度均值筛选后获得的图像区域分割出来,获得图 8(f)所示汉字编码点图像。

3.2 汉字编码点识别

从实际拍摄的图像中分割出来的汉字编码点图像(实拍图像)尺寸并不统一,本文用双线性插值法^[21]将分割出的每个实拍编码点图像的尺寸调整为 68×68 pixel,以适应识别网络的输入尺寸。将统一尺寸后的实拍图像进行高斯滤波处理,以降低实拍图像中噪声点的影响,然后用汉字编码点识别网络对其进行识别并输出实拍图像对应的汉字编码点身份号。图 9 所示即为汉字编码点识别网络的识别结果示例。

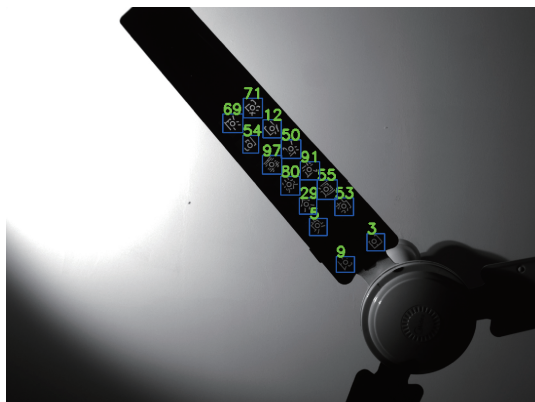


图 9 汉字编码点识别结果

Fig.9 Recognition result of Chinese character coded marks

3.3 汉字编码点中心定位

完成实拍图像中汉字编码点的分割及识别后,需要对这些编码点图像进行中心定位,本文提出的中心定位算法具体步骤如下。

1) 对分割得到的编码点图像进行二值化处理,并用 Canny^[20]边缘检测算子提取编码点轮廓;

2) 利用面积周长准则、轮廓层级准则对提取的编码点轮廓进行初步筛选,得到由待选的辅助圆轮廓与定位圆轮廓组成的轮廓组;

3) 对待选轮廓组中的两个轮廓分别进行最小二乘椭圆拟合,得到两个椭圆方程及相关参数;

4) 利用灰度验证准则确定拟合得到的两个椭圆分别为编码点辅助圆和定位圆;

5) 利用圆心距评价准则进行定位精度评价。

其中涉及的一系列准则及最小二乘椭圆拟合等具体内容如下:

(1) 面积周长准则:

计算由 Canny 算子提取的每个编码点轮廓的面积和周长,并设定轮廓的面积和周长筛选阈值以剔除不符合要求的轮廓;

(2) 轮廓层级准则:

检测所有轮廓的层级关系,对于内外嵌套的轮廓,外部轮廓为父轮廓,内部轮廓为子轮廓。如图 10 所示,汉字编码点定位圆为黑色小圆,其外还有一个半径稍大的同心白色小圆作为辅助圆,因此定位圆的轮廓存在父轮廓但不存在子轮廓,辅助圆的轮廓存在且仅存在一个子轮廓,利用轮廓层级准则筛选符合条件的一组或多组父子轮廓,剔除不符合轮廓层级准则的轮廓。

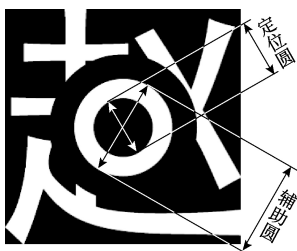


图 10 汉字编码点原图

Fig.10 Original picture of Chinese character coded mark

(3) 最小二乘椭圆拟合:

由于投影角度原因,汉字编码点成像后的辅助圆及定位圆通常为椭圆,因此本文使用最小二乘椭圆拟合方法对经过初步筛选后保留的待选轮廓组进行椭圆拟合。对于任意椭圆,其方程可表示为:

$$f(x, y) = x^2 + Axy + By^2 + Cx + Dy + E = 0 \quad (7)$$

对轮廓上数目为 $N(N \geq 6)$ 的点进行最小二乘椭圆拟合,即可求得椭圆参数 A, B, C, D, E ,由此确定拟合得到的椭圆方程,则该椭圆中心 (x_c, y_c) 可表示为:

$$\begin{cases} x_c = \frac{2BC - AD}{A^2 - 4B} \\ y_c = \frac{2D - AC}{A^2 - 4B} \end{cases} \quad (8)$$

椭圆的最小二乘拟合方法是寻找一个参数集合以确定椭圆方程,使该椭圆与待拟合坐标点之间的距离度量最小化,本文采用基于几何距离的椭圆拟合方法。对于椭圆轮廓上的每个点 p_i ,其几何距离定义为该点到椭圆最近点的距离,表示为:

$$d_{\min}^{(i)} = \sqrt{(x_i - x)^2 + (y_i - y)^2} \quad (9)$$

当椭圆拟合完成后,需对参与该椭圆拟合的每一个点进行几何距离验证,若某个点的几何距离超过设定的

距离阈值,则表示该点为椭圆的偏离点,当偏离点占轮廓点总数 N 的比值超过设定的比例阈值时,则表示该轮廓不符合椭圆拟合要求,剔除该轮廓,进行下一轮廓的最小二乘椭圆拟合。

(4) 灰度验证准则:

在对所有待选轮廓完成最小二乘椭圆拟合后,用灰度均值确认拟合得到的椭圆分别为辅助圆和定位圆。经轮廓层级准则筛选后的父子轮廓中,子轮廓为定位圆,父轮廓为辅助圆,实拍图像已二值化,定位圆成像椭圆的长短半轴 a_b, b_b 与辅助圆成像椭圆的长短半轴 a_w, b_w 均可从拟合的椭圆参数中得到,因此定位圆灰度均值 $Gray_{black}$ 与辅助圆灰度均值 $Gray_{white}$ 理论上可表示为:

$$\begin{cases} Gray_{black} = 0 \\ Gray_{white} = 255 \left(1 - \frac{a_b b_b}{a_w b_w} \right) \end{cases} \quad (10)$$

由于椭圆拟合可能存在误差,因此定位圆与辅助圆的实际灰度均值可能与理论值有所偏差,设定两圆的灰度均值阈值,若定位圆与辅助圆的灰度均值均在阈值范围内,则确认该组父子轮廓为所求辅助圆与定位圆的轮廓。图 11 所示为中心定位时拟合的定位圆和辅助圆。



图 11 汉字编码点定位圆及辅助圆

Fig.11 Positioning circle and auxiliary circle of Chinese character coded marks

(5) 圆心距评价准则:

确认拟合得到的椭圆分别为编码点的辅助圆和定位圆后,通过辅助圆圆心与定位圆圆心的像素距离来评价中心定位算法的精度。根据式(8)分别求出辅助圆圆心 O_w 的坐标 (x_w, y_w) 及定位圆圆心 O_b 的坐标 (x_b, y_b) ,则两圆心之间的像素距离 l_{bw} 可表示为:

$$l_{bw} = \sqrt{(x_b - x_w)^2 + (y_b - y_w)^2} \quad (11)$$

从汉字编码点设计角度讲,辅助圆和定位圆为同心圆,两者的圆心所处位置相同,其像素距离理论值为 0,因此实际计算得到的圆心像素距离 l_{bw} 越小,说明两圆圆心距离越接近理论值,中心定位算法的精度越高。设分割得到的编码点图像尺寸为 $l_{obj} \times l_{obj}$ pixel,本文认为当 $0 < l_{ap} < 0.003 l_{obj}$ 时,该组辅助圆与定位圆的椭圆拟合结果较好,中心定位精度较高。

4 实验与结果分析

本文提出的汉字编码点识别网络由卷积神经网络训

练得到,使用的深度学习框架是 **Keras**, 语言为 **Python**。为了验证本文提出的汉字编码点分割、识别及定位方法,本节设计了相关实验。实验场景如图 12 所示,其中 1 是由两台 Basler A102f 工业相机组成的双目立体视觉系统,其相机镜头型号为 Schneider Xenoplan 1. 4/17-0513,该系统用于获取实验中的实拍图像;2 是用于控制相机的计算机;3 是布设了汉字编码点的 LOGO 纸板,用于进行编码点分割实验;4 是布设了汉字编码点的风扇扇叶,主要用于进行汉字编码点的识别实验;5 是照明设备,用于调整光照强度。



图 12 实验场景
Fig.12 Experiment scene

4.1 编码点分割实验

为了验证汉字编码点的分割算法,本文从用于训练汉字编码点识别网络的 100 种汉字编码点中随机选出了 5 种汉字编码点并布设于带有复杂背景图案的 LOGO 纸板上,然后用双目立体视觉系统拍摄 LOGO 纸板在不同光照强度、不同空间位置等情况下的图像。LOGO 纸板与双目立体视觉系统的水平距离从 0.5 ~ 2.5 m 中取 5 个位置,各位置间相距约 0.5 m,每个位置都分别在偏暗、正常、偏亮 3 种不同的光照强度下拍摄图像,共 $5 \times 3 \times 2 = 30$ 幅 LOGO 纸板图像(双目立体视觉系统拍摄的每组图像包括左右共两幅图像)。图 13 所示为其中几幅不同光照强度和空间位置下的 LOGO 纸板图像。利用汉字编码点分割算法对 30 幅扇叶图像进行汉字编码点分割,最终对 $30 \times 5 = 150$ 个汉字编码点全部分割成功。图 14 所示为图 13 中成功分割汉字编码点后的图像,其中蓝色轮廓为 Canny 边缘检测算子提取的初步轮廓,绿色选框为最终分割的汉字编码点最小外接矩形。

4.2 识别实验及干扰模型

4.1 节中分割实验仅得到 150 个汉字编码点实拍图像,为了使识别实验的统计结果更具说服力,需获得更多的汉字编码点实拍图像。本文从用于训练汉字编码点识

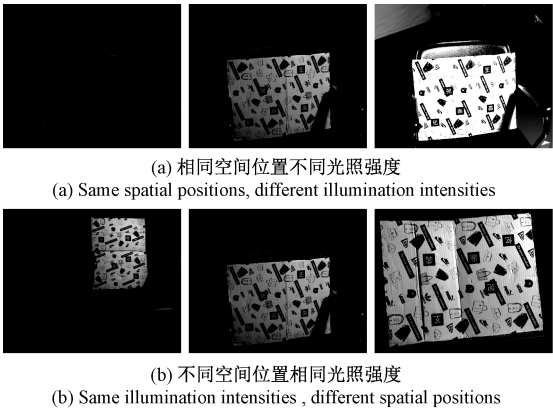


图 13 不同光照强度及空间位置的 LOGO 纸板图像
Fig.13 LOGO cardboard images for different illumination intensities and spatial positions

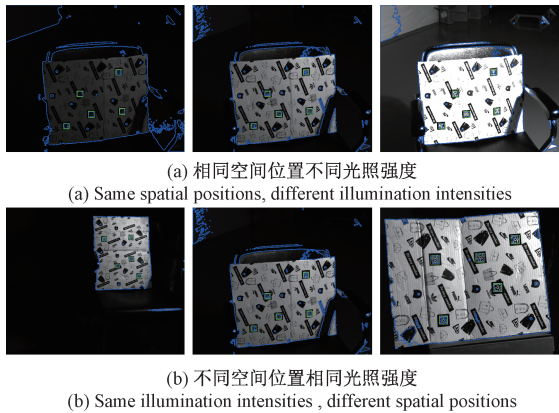


图 14 不同光照强度及空间位置的 LOGO 纸板图像汉字编码点分割

Fig.14 Chinese character coded mark segmentation of LOGO cardboard images for different illumination intensities and spatial positions

别网络的 100 种汉字编码点中随机选出了 14 种汉字编码点并布设于风扇扇叶上(见 3.2 节),然后用双目立体视觉系统拍摄风扇在不同光照强度、不同旋转位置等情况下的图像。扇叶旋转位置从一周 360° 中取 24 个不同位置,每个位置间的旋转角度大约为 15° ,每一个旋转位置都在偏暗、正常、偏亮 3 种不同的光照强度下拍摄 3 组图像,共 $24 \times 3 \times 2 = 144$ 幅扇叶图像。利用汉字编码点分割算法对所有扇叶图像进行汉字编码点分割,最终对 $144 \times 14 = 2\,016$ 个汉字编码点全部分割成功,其中光照强度为偏暗、正常、偏亮的汉字编码点实拍图像数量均为 672。用识别网络对不同光照强度的 3 组图像分别进行识别,识别率如表 1 所示。

由表 1 可知,对模拟图像识别率高达 98.3% 的汉字编码点识别网络在识别实拍图像时识别率并不理想,仅为 75.54%。对比汉字编码点的模拟图像与实拍图像后,发现实拍图像受到成像设备、光照等外界因素影响,图像

表 1 实拍编码点识别率(改进识别网络前)

Table 1 The real captured ceded mark recognition rate (before improving the identification network)

光照强度	识别量	识别错误量	识别率(%)
偏暗	672	232	65.47
正常	672	72	89.28
偏亮	672	189	71.87
综合	2016	493	75.54

中会存在噪声点,且各图像间对比度和亮度也不同,而虚拟相机法生成的模拟图像则过于理想化,没有噪声点且各图像间对比度和亮度均相同。

针对上述情况,本文提出一种干扰模型对识别网络的训练样本进行模拟干扰:1)对模拟图像添加各种噪声,包括椒盐噪声、条状噪声以及高斯噪声;2)随机调整模拟图像的对比度 α 与亮度 β ,其定义如下:

$$\begin{cases} \alpha = \frac{Gray_{max} - Gray_{min}}{255}, \alpha \in [0.3, 1] \\ \beta = Gray_{min}, \beta \in [0, 25] \end{cases} \quad (12)$$

其中, $Gray_{max}, Gray_{min}$ 分别为单张实拍图像中的最大和最小灰度值,对实拍图像进行灰度统计后可得对比度 α 与亮度 β 的范围,则模拟图像灰度值为 $Gray = 255\alpha + \beta$ 。图 15 所示为汉字编码点的实拍图像以及使用干扰模型添加噪声、对比度、亮度等影响因素前后的模拟图像之间的对比。

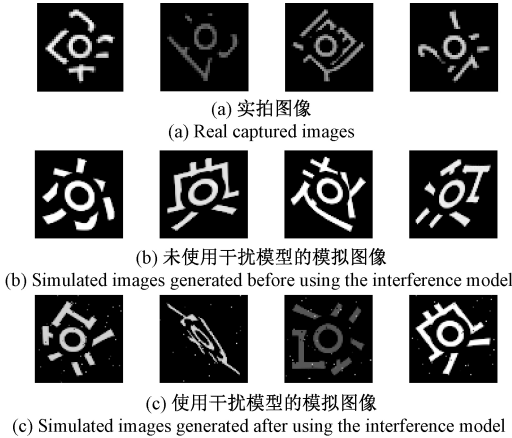


图 15 实拍图像与使用干扰模型前后生成的模拟图像对比

Fig.15 Comparison between real captured images and the simulated images generated before and after using the interference model

将训练样本进行改进后,对汉字编码点识别网络进行重新训练,训练过程中网络各参数仍按 2.4 节中设定,训练样本量为 16 万,训练周期为 50。为了更为直观地观察识别网络的训练效果,在每个训练周期结束后均使用改进后的模拟测试样本集对识别模型进行识别率测试。

图 16 所示为训练样本改进后汉字编码点识别网络的识别率与训练周期的关系,当训练周期为 30 时识别率曲线已基本趋于稳定,可见 50 个训练周期足以保证汉字编码点识别网络得到充分训练。

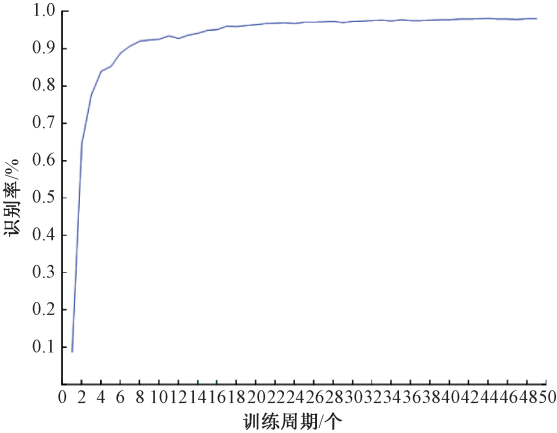


图 16 训练样本改进后网络识别率与训练周期关系

Fig.16 The relationship between recognition rate and training period after improving training samples

用实拍图像测试改进后的汉字编码点识别网络,结果如表 2 所示。由表 2 可知对训练样本进行改进后,汉字编码点对实拍图像的正确识别率可达 97.67%,对比改进之前有显著提升,且该识别网络完成每组 672 个汉字编码点识别的耗时均在 0.5 s 左右,平均每个汉字编码点约耗时 0.74 ms,因此利用识别网络识别汉字编码点效率很高。

表 2 实拍编码点识别率(改进识别网络后)

Table 2 The real captured ceded mark recognition rate (after improving the identification network)

光照强度	识别量	识别错误量	识别率/%
偏暗	672	24	96.43
正常	672	4	99.40
偏亮	672	19	97.17
综合	2 016	47	97.67

4.3 中心定位实验

完成汉字编码点识别实验后,对偏暗、正常、偏亮 3 组光照强度不同的编码点实拍图像分别进行中心定位实验。按照 3.3 节中步骤 1)~3)对汉字编码点的辅助圆和定位圆进行椭圆拟合,3 组共 2 016 个汉字编码点的辅助圆和定位圆全部拟合成功。用 3.3 节中的圆心距评价准则对每组 672 个汉字编码点的中心定位结果进行精度检验,结果如表 3 所示。

由表 3 可知正常光照下汉字编码点圆心距平均值有 0.128 pixel,光照强度偏亮或偏暗时其圆心距平均值有

表 3 不同光照强度下汉字编码点圆心距
Table 3 Circle center distance of Chinese character
code mark under different light intensities (pixel)

光照强度	最大圆心距	最小圆心距	平均圆心距	标准差
偏暗	0.209	0.037	0.154	0.027
正常	0.173	0.032	0.128	0.021
偏亮	0.182	0.024	0.137	0.023

所下降,但均在 0.2 pixel 以内。进行实验的汉字编码点实拍图像尺寸约为 68×68 pixel,根据圆心距评价标准(详见本文 3.3 节),当圆心距 $l_{bw} < 0.003 \times 68 = 0.204$ pixel 时,即可认为中心定位结果精度较高,因此该实验结果表明本文提出的汉字编码点中心定位算法精度较高,且受光照强度的影响较小。

5 结 论

本文提出了一种以汉字作为编码特征的编码标记点,并利用基于卷积神经网络的汉字编码点识别网络对其进行识别,结合本文提出的汉字编码点分割算法以及中心定位算法,可准确从图像中分割、识别汉字编码点,并对其进行中心定位。该方法的优点有:1)用汉字编码点识别网络替代传统识别算法,解除了识别算法对编码复杂度的限制,任何具有唯一性的特征都可用于编码点的编码设计,仅汉字作为编码就有上万种编码种类,极大扩展了编码空间,保证了足够多的编码点种类数量。2)利用虚拟相机法并结合可添加噪声、对比度等影响的干扰模型,可方便地生成足量且优质的训练样本用于训练识别网络,保证识别网络训练充分。3)汉字编码点识别网络对实际拍摄的汉字编码点的正确识别率达 97.96%,且识别效率高,平均每个汉字编码点识别所耗时间约 0.74 ms,网络泛化性好,受光照等因素影响小。4)汉字编码点分割算法鲁棒性强,中心定位算法精度高。

本文提出的汉字编码点在一定模糊效应下,仍能够保留其基本架构特征,具有一定抗模糊性能,并通过卷积神经网络对其进行识别,该方法未来或可应用于研究离焦模糊或运动模糊等项目中。

参考文献

[1] 李晓峰,张瑜,骆念武. 摄影测量中环状编码点的检测方法[J]. 现代制造工程, 2012(2): 118-121.
LI X F, ZHANG Y, LUO N W. Detection method of ring coding points in photogrammetry [J]. Modern Manufacturing Engineering, 2012 (2): 118-121.
[2] 马扬飏,钟约先,郑聆,等. 三维数据拼接中编码标志点的设计与检测[J]. 清华大学学报(自然科学

版), 2006, 46(2): 169-171.
MA Y B, ZHONG Y X, ZHENG L, et al. Design and detection of coding markers in three-dimensional data mosaic [J]. Journal of Tsinghua University (Natural Science Edition), 2006, 46 (2): 169-171.
[3] SHORTIS M R, SEAGER J W, ROBSON S, et al. Automatic recognition of coded targets based on a Hough transform and segment matching [C]. Videometrics VII. International Society for Optics and Photonics, 2003, 5013: 202-209.
[4] AHN S J, RAUH W, KIM S I. Circular coded target for automation of optical 3D-measurement and camera calibration [J]. International Journal of Pattern Recognition and Artificial Intelligence, 2001, 15 (6): 905-919.
[5] LI S Z, MADDAH-ALI M A, AVESTIMEHR A S. Coding for distributed fog computing [J]. IEEE Communications Magazine, 2017, 55(4):34-40.
[6] WANG Y W, CAI B L, WANG K Y, et al. Out-of-focus color camera calibration with one normal-sized color-coded pattern [J]. Optics & Lasers in Engineering, 2017, 98:17-22.
[7] 王文俊,殷曦敏. 手机二维码识别技术在大型仪器设备管理中应用[J]. 实验室研究与探索, 2015, 34(5): 278-281.
WANG W J, YIN X M. Application of mobile phone two-dimensional code recognition technology in management of large-scale instruments and equipment [J]. Laboratory Research and Exploration, 2015, 34(5): 278-281.
[8] 余博,田福庆,梁伟阁. 基于深度卷积变分自编码网络的故障诊断方法[J]. 仪器仪表学报, 2018, 39(10):27-35.
SHE B, TIAN F Q, LIANG W G. Fault diagnosis method based on deep convolution variational self-coding network [J]. Journal of Instruments and Instruments, 2018,39(10): 27-35.
[9] 李绣心,凌志刚,邹文. 基于卷积神经网络的半监督高光谱图像分类[J]. 电子测量与仪器学报, 2018, 32(10): 95-102.
LI X X, LING ZH G, ZOU W. Semi-supervised hyperspectral image classification based on convolution neural network [J]. Journal of Electronic Measurement and Instruments, 2018, 32 (10): 95-102.
[10] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
[11] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich

- feature hierarchies for accurate object detection and semantic segmentation [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014: 580-587.
- [12] 熊有伦. 机器人技术基础[M]. 武汉: 华中科技大学出版社, 1996: 15-29.
XIONG Y L. Basic robot technology [M]. Wuhan: Huazhong University of Science and Technology Press, 1996: 15-29.
- [13] NAIR V, HINTON G E. Rectified linear units improve restricted boltzmann machines [C]. International Conference on Machine Learning (ICML 2010), 2010: 807-814.
- [14] WU H B, GU X D. Max-pooling dropout for regularization of convolutional neural networks [C]. International Conference on Neural Information Processing. Springer, Cham, 2015: 46-54.
- [15] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: A simple way to prevent neural networks from overfitting [J]. The Journal of Machine Learning Research, 2014, 15(1): 1929-1958.
- [16] RAZAVIAN A S, AZIZPOUR H, SULLIVAN J, et al. CNN features off-the-shelf: an astounding baseline for recognition [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2014: 512-519.
- [17] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]. Proceedings of the IEEE International Conference on Computer Vision. 2017: 2999-3007.
- [18] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. Nature, 2015, 521: 436-444.
- [19] LECUN Y A, BOTTOU L, ORR G B, et al. Efficient backprop [C]. Neural networks: Tricks of the Trade, 2012: 9-50.
- [20] CANNY J. A computational approach to edge detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, PAMI-8(6): 679-698.
- [21] ZHOU R G, HU W W, FAN P, et al. Quantum realization of the bilinear interpolation method for NEQR[J]. Scientific Reports, 2017, 7(1): 2511.

作者简介



陶聪, 在 2016 年于南京航空航天大学获得学士学位, 现为南京航空航天大学硕士研究生, 主要研究方向为机器视觉。

E-mail: taocccong@163.com

Tao Cong received his B. Sc. from Nanjing University of Aeronautics and Astronautics in 2016. Now, he is a master student in Nanjing University of Aeronautics and Astronautics. His main research interest is machine vision.



张丽艳 (通信作者), 2001 年于南京航空航天大学获得博士学位, 现为南京航空航天大学教授、博士生导师, 主要研究方向为三维视觉检测、智能制造、逆向工程。

E-mail: zhangly@nuaa.edu.cn

Zhang Liyan (Corresponding author) received her Ph. D. degree from Nanjing University of Aeronautics and Astronautics in 2001. She is now a professor and Ph. D. supervisor at Nanjing University of Aeronautics and Astronautics. Her main research interests include 3D vision inspection, intelligent manufacturing, and reverse engineering.